IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 4, April 2015.
www.ijiset.com
ISSN 2348 – 7968

Hinder Blackhat's Anonymous Threat by identifying user's Non meticulous Traversal Behavior Patterns

P.Kavitha[1]
Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, Tamil Nadu, India.

Dr.G.N.K.Suresh babu[2]
Associate Professor, GKM College of Engineering and Technology, Chennai, Tamil Nadu, India.

K.Sankar[3]
Research Scholar, Department of Computer Science, Bharathiar University, Coimbatore, Tamil Nadu, India.

Abstract:

Organizations prolong to slot in the Internet as a key factor of their operations, which in turn increase the risk of cyber-threat level inside and outside an organization. The proposed system involves in identifying insider attack by identifying frequent mining pattern. Although many studies have focused on these issues with certain techniques, it fails to satisfy an organization expecting security level. Most of the technique uses simply the mining pattern, which needs maximum support to identify the mining behavior of the user to predict his abnormal behavior. Those techniques also increase the overhead in identifying the abnormal behavior pattern. To resolve this problem, the proposed system uses a novel approach which involves in monitoring both frequent mining pattern and subsequent database queries of individual user to identify the weighing factor using from which the proposed system can able to detect insider attack with less overhead and with minimum support.

Keywords: Behavior patterns, Cyber-threat, Insider attack, Mining pattern.

1. Introduction:

In recent years, widespread adoption of the internet has resulted into rapid advancement in information technologies. The internet is used by the general population for the purposes such as financial transactions, educational activities and countless other activities. This growth of the Internet use has unfortunately been accompanied by a growth of malicious activity in the Internet and also in Intranet. There are many ways in doing such malicious activities; one of the most notorious types is insider attack. Insider attacks can have an effect on all computer security elements and range from stealing sensitive data to injecting Trojan viruses in a system or network. Insiders also may affect system availability by overloading computer/network storage or processing capacity, leading to system crashes. Internal intrusion detection systems .

(IDS) protect organizations against insider attacks, but deploying such systems is not easy. Rules must be established to ensure that unintended attack warnings are not triggered by employees. In 2008, a noteworthy insider attack occurred when Terry Childs, a network engineer for the San Francisco Department of Telecommunications and Information Services, altered the city's network passwords, locking FiberWAN access for 12 days. Childs was found guilty of felony network tampering. The work required to regain system control cost the city of San Francisco $900,000, and 60 percent of city services were affected by the insider attack [1].

Many companies focus all of their security efforts on keeping out hackers and other network intruders. But from my perspective, the threat posed by an insider attack is actually greater than that of external hackers and viruses. It is known that the full attack path of an external hacker, the first step is to gain internal access. Organizations expend an extraordinary amount of resources on protecting their perimeter specifically to counter this threat. For the malicious insiders, though, these countermeasures don't apply, since these people are already on the inside and they enjoy a certain implicit trust.

The proposed system uses the combination of indentifying the behavior pattern of an individual and his/her accessibility permission over the database or the network. Many studies related to traversal path prediction and web pre fetching thus emerge one after another [2][3]; the mining techniques for frequent path traversal patterns [2] are also taken into account too for effectively digging up useful information within the huge amount of data. Section 2 covers the literature survey; proposed system along

ISSN 2348 – 7968

with the detailed description of insider attack is covered in section 3. Performance analysis is done in section 4. Conclusion and future work is explained in section 5.

2. Literature survey:

An insider attack is a legitimate user within an organization or an environment, who has been granted every access to systems and information resources, but whose procedures are always contradict to policy, and he/she will always play a role in accessing or stealing information with their freedom over the organization [7].

B. D. Davison et al [8] proposes Command line calls issued by users, System call monitoring for unusual application events, Organization policy management rules and compliance logs. The type of analysis used is primarily the modeling of statistical features, such as the frequency of events, the duration of events, the co-occurrence of multiple events combined through logical operators, and the sequence or transition of events. However, most of this work failed to reveal or clarify the user's intent when issuing commands.

Szymanski and Zhang [9] proposed recursively mining the sequence of commands by finding frequent patterns, encoding them with unique symbols, and rewriting the sequence using this new coding. Though this system seems to be promising it cannot easily implemented in a real-world setting.

James Newsome et al proposes Signature Generation Algorithm which defines polymorphic signature generation problem which proposes classes of signature suited for matching polymorphic worm payloads and generate signatures in these classes. Since it uses the history of previous attacks, it cannot detect new attacks or previously unseen attacks [4].

Christopher Kruegel et al [5] presents an Intrusion detection system that uses a number of anomaly detection techniques to detect attacks against web servers. Behavior models are built by performing a statistical analysis on historical data. They detect deviations from the learned patterns of the user behavior. It can detect new attacks but it produces false alarm for legitimate but previously unseen system behavior.

Yih Huang et al [6] proposes a general-purpose framework that harnesses the power of lightweight virtualization to track applications' interactions in a scalable and efficient manner. It involves in storing huge data set for long time, since it requires large memory and also it consumes time.

The user profiling idea for insider threat detection in particular and anomaly detection in general, is certainly involves in consuming time for finding the factor to detect the anomaly [10]. Thus the proposed system uses SPADE algorithm for finding users frequent mining pattern along with time, so that the mining pattern for the user can be formed efficiently with little time consumption.

3. Proposed system

Proposed system uses predefined promising knowledge like knowing the user's actual work that he should perform and compare it with his regular frequent transaction to identify an attack, through which it will be possible to strengthen the security. Communications are categorized as sessions which identify the mapping between web server request and subsequent DB queries.

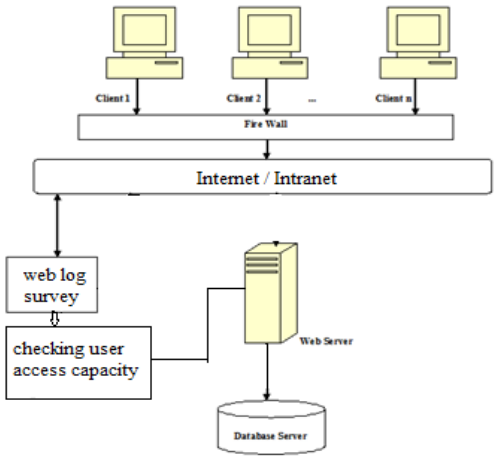

Fig 1. System Architecture

Using this approach, at database side, we are able to tell which DB transaction corresponds to which client request. This helps us to identify the mapping between web server request and corresponding Database queries. By using this mapping model, we detect abnormal behavior on a session by checking it with the predefined known factor about the user.

Fig 1 explains about the overall architecture design of the proposed system. All users mining pattern connected to the internet/intranet within in an organization will be monitored for identifying the user behavior along with his/her utmost permission granted to access, by knowing this a determining factor will be calculated for identifying the attack.

3.1. SPADE Algorithm for identifying frequent pattern:

The proposed system uses SPADE for predicting the frequent mining pattern of the user. Sequence procedure enables proposed system to

ISSN 2348 – 7968

perform sequence discovery. It is similar to that of association rule, but sequence discovery goes one step further than association discovery by taking into account the ordering or timing of the relationship among items. By knowing the timing factor the proposed system will not only track the users actual frequent mining pattern, but also the time he spends each session.

3.2. Web Log survey to predict user frequent 3.2.1 transaction pattern – using SPADE algorithm:

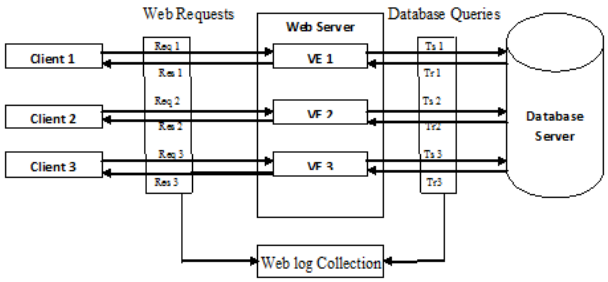

Fig 2. Web Log Collection for Identifying User Frequent Mining Pattern

Fig 2 explains about the collection of weblogs of each and every user in an organization connected to the internet or intranet. Both the web request and the database queries of the users along with the time factor using SPADE will be collected and segregate as in table 1.

USER LOG FILES		
SID	**TIME**	**MINING TERMS**
1	**10**	**O P**
1	**15**	**M N O**
1	**20**	**M N R**
1	**25**	**M O P R**
2	**15**	**M N R**
2	**20**	**Q**
3	**10**	**M N R**
4	**10**	**P S T**
4	**20**	**N R**
4	**25**	**M S T**

Table 1: Log files of users frequent mining pattern

The terms M, N, O, P, Q, R, S, T in table 1 and table 2 are referred to the mining terms that the user visit every time from his/her node.

Consider table 1 which has the sequences of mining pattern of an individual at an organization, form the frequent sequences is obtained. Table 2 shows the frequent and possible pattern about which the user do as his regular activities about viewing a page or accessing an item.

FREQUENT SEQUENCES		
Frequent 1 - Sequences		**Time**
M	**4**	**15**
N	**4**	**15**
P	**2**	**10**
R	**4**	**15**
Frequent 2 - Sequences		
MN	**3**	**15**
MR	**3**	**20**
N->M	**2**	**15**
NR	**4**	**20**
P->M	**2**	**25**
P->N	**2**	**10**
P->R	**2**	**10**
R->M	**2**	**10**
Frequent 3 - Sequences		
MNR	**3**	**20**
NR->M	**2**	**15**
P->NR	**2**	**20**
P->N->M	**2**	**20**
P->R->M	**2**	**10**
Frequent 4 - Sequences		
P->NR->M	**2**	**20**

Table 2: Identification of Frequent pattern using time factor

This will be checked against the user's utmost permission to access data within an organization. By that the user abnormal behavior can be easily identified if the time and the access pattern get deviates from the regular patterns.

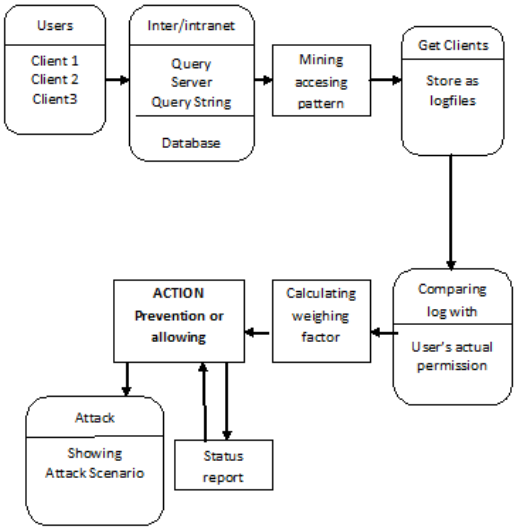

Fig 3. Data flow diagram for identifying attack based on mining pattern

ISSN 2348 – 7968

Fig 3 explains about the flow of proposed system in identifying whether there present attack or not. The status report and the attack scenario can be easily identified with the proposed system with less overhead compare to that of existing system. Fig 3 explains about the identifying of attack based on user's frequent mining behavior.

4. Performance analysis:

Each Session will have some set of requests and queries from user, from which the accessing pattern will be identified, and then it will be compared against with the user's actual work which is predefined along with his legitimate actual pattern of accessing. The experimental results were done several times by varying the accessing pattern along with various time factors. The proposed system shows efficient result with more accuracy. The system performance has been analyzed by varying the mining pattern of user and time factor regularly and the performance is shown in Fig 4.

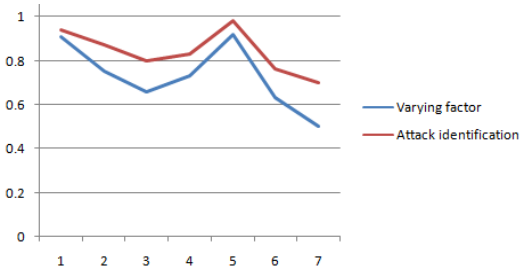

Fig 4: Experimental result

The above experimental result shows that our proposed method shows high accuracy in predicting insider attack within an organization.

5. Conclusion and Future work:

The proposed work uses both the mining pattern along with time factor, which helps to identify the mapping between web server request and subsequent DB servers of any user. Through this approach the system can identify the transaction corresponds to client request and can check it against with their regular mining pattern. Using this mapping model, the proposed system can detect abnormal behavior on a session at client level. Thus it can detect any kind of unknown anonymous access of the network. The system performance can be improved more by adding some heuristic approaches as in Honeypot.

Reference:

[1]http://www.techopedia.com/definition/26217/insider-attack

[2] M.S. Chen, J.S. Park, and P.S. Yu,: "Efficient Data Mining for Path Traversal Patterns", IEEE Transactions on Knowledge and Data Engineering, 209--220(1998)

[3] S. Vallamkondu, and L. Gruenwald,: "Integrating Purchase Patterns and Traversal Patterns to Predict HTTP Requests in E-Commerce Sites", Proc. of IEEE Int. Conf. on E-Commerce(2003)

[4] James Newsome, Brad Karp and Dawn Song, "Polygraph: Automatically Generating Signatures for Polymorphic Worms," Proceedings of the Security and Privacy, IEEE Symposium on May 2005.

[5] Christopher Kruegel and Giovanni Vigna, "Anomaly Detection of Web Based Attacks," Association for Computing Machinery Conference. Computer and Comm. Security (CCS '03), Oct. 2003.

[6] Yih Huang, Angelos Stavrou, Aup K.Ghosh and Sushil Jajodia, "Efficiently Tracking Application Interactions Using Lightweight Virtualization," Proceedings of the First Workshop on Association for Computing Machinery, Oct. 2008.

[7] Mark Maybury, Penny Chase, Brant Cheikes, Dick Brackney, Sara Matzner, Tom Hetherington, Brad Wood, Conner Sibley, Jack Marin, and Tom Longsta . Analysis and detection of malicious insiders. In Proceedings of the International Conference on Intelligence Analysis, 2005.

[8] B. D. Davison and H. Hirsh. Predicting sequences of user actions. In Working Notes of the Joint Workshop on Predicting the Future: AI Approaches to Time Series Analysis, 15th National Conference on Articial Intelligence/15th International conference on Machine Learning , pages 512. AAAI Press, 1998.

[9] Boleslaw K. Szymanski and Yongqiang Zhang. Recursive data mining for masquerade detection and author identification. In Proceedings of the 13rd Annual IEEE Information Assurance Workshop . IEEE Computer Society Press, 2004.

[10] Schonlau, M., DuMouchel, W., Ju, W., Karr, A., Theus, M., Vardi, Y.: Computer intrusion: Detecting masquerades. Statistical Science 16(1), 58–74 (2001).